



**Réalité virtuelle : Simulation sonore en trois dimensions
par Nicolas Tsingos. Phobies à l'épreuve du virtuel,
entretien avec Isabelle Viaud-Delmon ; propos recueillis
par Dominique Chouchan. Le jeu vidéo est un produit
de très haute technologie par David Alloza**

Nicolas Tsingos, Isabelle Viaud-Delmon, David Alloza

► **To cite this version:**

Nicolas Tsingos, Isabelle Viaud-Delmon, David Alloza. Réalité virtuelle : Simulation sonore en trois dimensions par Nicolas Tsingos. Phobies à l'épreuve du virtuel, entretien avec Isabelle Viaud-Delmon ; propos recueillis par Dominique Chouchan. Le jeu vidéo est un produit de très haute technologie par David Alloza. Les Cahiers de l'INRIA - La Recherche, 2009, La guerre des trous noirs, 427 février 2009. inria-00527609

HAL Id: inria-00527609

<https://inria.hal.science/inria-00527609>

Submitted on 19 Oct 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RÉALITÉ VIRTUELLE

Simulation sonore en trois dimensions

En synthèse d'images, on s'intéresse à la simulation des formes, des couleurs, des ombres, de l'éclairage... La sonorisation d'un monde synthétique est tout aussi exigeante si l'on veut qu'elle reflète notre perception spatiale des sons.

Nous savons localiser un événement sonore sans le voir, que ce soit une collision automobile, le passage d'un train, un orchestre de rue à proximité..., du moins si notre système auditif fonctionne correctement. Celui-ci nous permet en effet de nous faire une représentation spatiale des sources sonores qui nous environnent. Si l'on veut simuler sur ordinateur une scène dans laquelle se produit une collision, juste après le passage d'un train et pendant qu'un orchestre joue non loin de là, la scène simulée ne sera crédible qu'à prix d'un « rendu sonore » réaliste, c'est-à-dire de la restitution de cette représentation spatiale. Une telle simulation est toutefois très gourmande en puissance de calcul. Comment l'optimiser ? C'est l'objet de la technologie Audile* que nous avons mise au point et qui exploite les connaissances sur la perception humaine^(1,2).

D'une manière générale, une écoute spatialisée, avec un casque stéréophonique ou un système multi-haut-parleurs, suppose que soient reproduites les caractéristiques acoustiques telles qu'elles auraient été induites par des sources sonores réelles. Or de la même manière que la vision en relief dépend des images, non identiques, reçues par chaque œil, la percep-

tion sonore tridimensionnelle dépend des sons, légèrement différents, reçus par chacune des oreilles. La simulation doit notamment tenir compte du fait que l'onde de pression émise par une source sonore ne les atteint pas exactement au même instant, sauf si la source est équidistante des deux oreilles.

Autre phénomène à considérer : les légères différences d'amplitude entre les sons perçus par l'une et l'autre oreille, du fait de la diffraction des ondes sonores au niveau du visage, du buste et du pavillon des oreilles. En effet, la diffraction n'est pas équivalente de part et d'autre de la tête : elle dépend de la provenance des ondes et de leur direction de propagation. S'ajoutent à cela les effets de réverbération des sons*, qui jouent un rôle essentiel dans la perception de la distance des sources et de leur environnement. Ils doivent donc aussi être simulés.

Alors seulement l'auditeur immergé dans un environnement de synthèse sera en mesure de percevoir l'univers sonore en trois dimensions (fig. 1). Le problème est que dans une scène réelle, de très nombreuses sources interviennent. Jusqu'à présent, il n'était possible d'en simuler qu'une infime partie. Du coup, l'habillage audio, celui des jeux vidéo par exemple, se réduisait souvent à une musique d'ambiance, éventuellement complétée par quelques sons se produisant sur le devant de la scène simulée. En théorie, des bruits comme ceux de l'écoulement d'une



Fig. 1 : La spatialisation des sons est essentielle pour l'immersion dans les jeux vidéo. Le rendu sonore de l'incendie montré sur ce cliché (jeu *Alone in the dark*) a été possible en utilisant la technologie Audile.

* Audile est une librairie logicielle développée par l'équipe Reves de l'INRIA Sophia Antipolis - Méditerranée depuis 2003 (voir http://interstices.info/jcms/c_30410/simulation-de-scenes-sonores-spatialisees-complexes). Elle permet de mettre en œuvre un ensemble de techniques pour le rendu audio-perceptif. Ces technologies sont brevetées depuis avril 2008.

* La réverbération d'un son désigne l'effet sonore résultant de la réflexion de l'onde de pression sur les solides (sols, murs...) qu'elle rencontre lors de sa propagation.

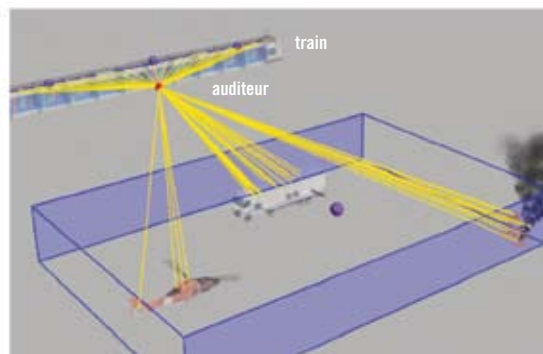
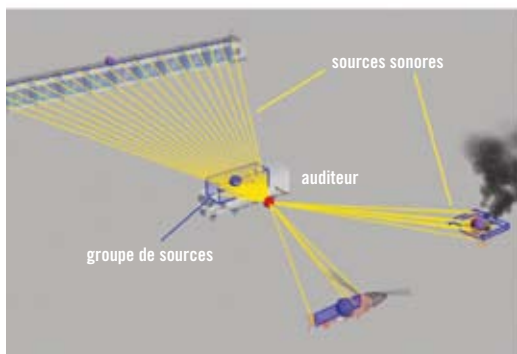


Fig 2. : Le regroupement des sources sonores s'adapte dynamiquement à la position de l'auditeur (boule rouge) et à l'importance relative des signaux sonores qu'il perçoit. Les nombreuses sources (lignes jaunes) sont regroupées en quatre sources cibles (boules violettes). Dans la configuration de gauche, le système a affecté une source unique par objet. Dans la configuration de droite, l'utilisateur s'est rapproché du train qui se voit alors représenté par trois sources, les autres objets se regroupant en une source unique.

rivière, d'un mouvement de foule, des impacts d'une explosion... nécessiteraient de gérer des centaines de sources sonores.

D'où l'idée de tirer profit de nos connaissances sur l'humain. Car malgré son acuité, notre audition a elle aussi ses limites : en nombre de sons susceptibles d'être perçus simultanément, en résolution spatiale... Notre cerveau fait en permanence des compromis pour pouvoir interpréter la complexité des univers sonores dans lesquels nous nous trouvons. C'est sur la mise en œuvre de tels compromis que se fonde l'outil logiciel Audile.

Exemple de limite de notre système auditif : un son intense (une violente explosion) peut nous rendre inaudible un son plus faible (la fermeture d'une porte). On parle de masquage acoustique. Les algorithmes mis en œuvre dans Audile exploitent ce phénomène et permettent de déterminer les sources audibles, à chaque instant de la simulation, sur la base d'une connaissance *a priori* de leur niveau sonore et de leurs fréquences (ou spectre sonore). De la même manière qu'on éliminerait les zones cachées d'une image tridimensionnelle, les sources inaudibles sont éliminées, ce qui libère de précieuses ressources de calcul.

Un autre exemple d'application de connaissances acoustiques concerne les limitations, dans le monde réel, de notre perception spatiale des sons : nous avons tous fait l'expérience de l'impossibilité de distinguer des voix ou de la musique au sein d'une atmosphère très bruyante. L'idée est alors de regrouper les sources proches les unes des autres en une source unique (fig. 2). Dans un premier temps, on en fait la somme, puis cette somme est traitée en tenant compte des indices acoustiques relatifs à la localisation de chaque composante du groupe (fig. 2). Si l'on se déplace dans une scène virtuelle, ce regroupement des sources est réajusté en temps réel en fonction de la position et de l'importance relative de chacune d'entre elles.

Un avantage majeur de ces algorithmes est de pouvoir adapter les traitements du son aux

capacités de calcul spécifiques de chaque plateforme (ordinateurs, consoles de jeux...). Divers secteurs industriels s'y intéressent, en particulier ceux de la simulation interactive (simulateurs de conduite...) et du jeu vidéo (fig. 1). La technologie Audile⁽³⁾ est par exemple mise en œuvre dans deux jeux du studio français Eden Games / Atari : *Test Drive Unlimited* (jeu de course automobile) et *Alone in the Dark: Near Death Investigation* (qui plonge le joueur dans un monde d'horreur). Un prototype destiné aux communications vocales spatialisées, pour des jeux multi-joueurs distribués, a également été réalisé en 2006, dans le cadre du projet « Optimisation perceptive du rendu audio » (Opera) du Réseau national des technologies logicielles (RNTL), en collaboration avec les laboratoires français Orange Labs et Virtools (filiale de Dassault Systemes)⁽⁴⁾.

Mais Audile intéresse également le monde de l'urbanisme et de l'architecture : les architectes du projet de tramway, à Nice, s'en sont servis pour évaluer une maquette interactive de certains aménagements. Enfin, au-delà de l'ingénierie, ce type de simulation sonore commence à trouver des applications thérapeutiques (voir l'entretien avec Isabelle Viaud-Delmon), par exemple pour le traitement de phobies. La réalité virtuelle pourrait en effet contribuer à corriger des désordres psychologiques liés à des problèmes de perception visuelle et auditive.

Nicolas Tsingos, chercheur INRIA au sein de l'équipe-projet Reves, travaille actuellement sur la synthèse et la spatialisation des sons, ainsi que sur les interactions sons-images pour les mondes virtuels interactifs.

⁽¹⁾ Nicolas Tsingos, Emmanuel Gallo and George Drettakis. *Perceptual Audio Rendering of Complex Virtual Environments*, SIGGRAPH 2004.

⁽²⁾ Thomas Mœck et al., *Progressive Perceptual Audio Rendering of Complex Scenes*. I3D, Symposium on Interactive 3D Graphics and Games, 2007.

⁽³⁾ voir : <http://www-sop.inria.fr/reves/audile/>

⁽⁴⁾ voir : <http://www-sop.inria.fr/reves/OPERA/>



D.R.

Agoraphobie, claustrophobie, etc. autant de troubles parfois très handicapants que l'immersion dans des mondes virtuels pourrait permettre de mieux maîtriser.

De la peur de l'enfermement à la terreur des foules, quels sont les points communs ?

Isabelle Viaud-Delmon : Il semble que ces phobies regroupent toutes des problèmes liés à un dysfonctionnement de l'intégration multisensorielle. En principe, nous percevons le monde extérieur de manière unifiée, alors que notre

impact. Curieusement, nous avons par exemple observé que les personnes anxieuses mettaient en place des mécanismes d'adaptation plus efficaces que les non anxieuses. Mais surtout, les résultats de cette étude, réalisée à la fin des années 1990, nous ont incités à poursuivre ces expériences et à tester leur intérêt dans un cadre clinique.

Avez-vous observé une amélioration de l'état des patients phobiques ?

I. V.-D. : Au centre « Emotions » de l'hôpital de la Pitié-Salpêtrière, nous avons travaillé sur plusieurs cohortes de patients. Ainsi, dans le cadre de la thèse de Fervel Znaïdi, un dispositif de réalité virtuelle a été expérimenté sur des sujets souffrant de phobies liées à l'espace (phobies des grands espaces, des hauteurs, de l'enfermement...). Après une dizaine de séances d'immersion dans des scènes virtuelles auxquelles ils devaient faire face, leur état s'est nettement amélioré.

D'où la nécessité, pour recréer des univers réalistes, de doter les environnements virtuels d'une dimension sonore ?

I. V.-D. : A ce jour, les sons sont peu présents dans les mondes virtuels ou, s'ils le sont, ils sont rarement restitués avec les indices acoustiques qui

Isabelle Viaud-Delmon, chercheur CNRS, est neuropsychologue. Elle vient de rejoindre l'unité mixte CNRS-Ircam de Sciences et technologies de la musique et du son (STMS) après avoir exercé à l'hôpital de la Pitié-Salpêtrière.

Entretien avec Isabelle Viaud-Delmon Phobies à l'épreuve du virtuel

perception résulte de la synthèse d'informations provenant d'une multitude de capteurs hétérogènes : capteurs chimiques pour la vision, capteurs de pression pour le toucher, accéléromètres pour le mouvement propre. Or certains patients éprouvent des difficultés à mettre en relation ces différentes informations. Ils se plaignent entre autres de vertiges ou de malaises, bien que leurs systèmes visuel, vestibulaire notamment, pris individuellement fonctionnent parfaitement. Le mal des transports, par exemple, peut relever de ce type de « conflit » sensoriel, comme l'angoisse au sein d'un ascenseur.

D'où vous est venue l'idée d'utiliser la réalité virtuelle ?

I. V.-D. : Elle nous est apparue comme une évidence, à l'époque où je travaillais sous la direction d'Alain Berthoz, au sein du Laboratoire de physiologie de la perception et de l'action (Collège de France). Avec un système de réalité virtuelle, on peut en effet créer à volonté des situations de conflit sensoriel puis étudier leur

permettraient de les localiser. Or l'audition offre une perception de l'espace bien plus complète que la vue (on ne voit pas derrière soi par exemple). Le son favorise une immersion de meilleure qualité, notamment au plan émotionnel. Cela suppose de synthétiser un monde sonore capable de reproduire les caractéristiques spatiales des sons, d'où l'intérêt d'outils permettant de contrôler la spatialisation du son en temps réel. Nous avons ainsi travaillé sur la phobie des chiens, dans le cadre du projet européen Crossmod*, afin d'étudier la sensibilité des personnes phobiques aux décalages entre stimuli auditifs (dans ce cas les aboiements) et visuels (image du chien) : en les immergeant dans un environnement virtuel, on peut manipuler (faire varier) ces décalages donc analyser leur caractère anxigène. Nous travaillons également sur le problème des acouphènes de manière à aider le patient à en maîtriser la perception. Un essai clinique à ce sujet vient de démarrer en collaboration avec l'hôpital européen Georges Pompidou.

Propos recueillis par Dominique Chouchan

* Le projet Crossmod (2005-2008) portait sur les techniques de rendu perceptif en réalité virtuelle (www.crossmod.org).



EDEN GAMES / ATARI

La sophistication croissante des jeux vidéo est à la mesure de l'ambition de leurs concepteurs qui les veulent de plus en plus réalistes, au plan sonore autant que visuel.



D.R.

David Alloza occupe un poste de directeur technique au sein de la société française Eden Games (filiale du groupe Atari).

«Notre activité consiste pour l'essentiel à concevoir et à développer des jeux vidéo sur des consoles de jeu de type PlayStation3* et Xbox360*. Nous utilisons à cet effet des kits de développement fournis par les constructeurs. En ce qui concerne notre intérêt pour les travaux de l'INRIA, il fait notamment suite à une publication présentée au Siggraph*, en 2004, sur des techniques de spatialisation du son dans un monde virtuel [voir l'article de Nicolas Tsingos]. La problématique initiale des chercheurs était toutefois très différente de la nôtre. Il s'agissait pour eux de restituer un environnement sonore le plus réaliste possible

La majorité des opérations de rendu audio repose désormais sur le software. En revanche, la limite qui en découle concerne le temps de calcul. Notre objectif était de réduire à environ 15 % la part de la simulation audio dans le volume global de calcul, tout en visant une très bonne qualité de rendu sonore. A ce titre, les solutions proposées par les chercheurs de l'INRIA, qui exploitent certaines connaissances acoustiques, sont précieuses. Elles permettent d'éviter des calculs inutiles, en particulier les calculs de sons *a priori* inaudibles (car masqués par d'autres sons par exemple). C'est cette approche que nous avons mise en œuvre dans nos deux derniers jeux vidéo (*Test Drive Unlimited* et *Alone in the Dark*), en développant un code adapté à nos contraintes industrielles. Pour *Test Drive Unlimited* (voir les captures d'écran), un jeu multijoueur de course automobile, nous voulions restituer l'environnement sonore complet des véhicules des différents joueurs: sons de roulement des pneus et de transmission dans le moteur, sons liés à l'aérodynamique (sifflement de l'air sur le pare-brise...), etc. Non seulement cela devenait possible, mais nous avons ainsi réussi à réduire des deux tiers la charge de calcul consacrée au son. Dans le jeu *Alone in the Dark*, la complexité est d'une autre nature. Il s'y produit par exemple un grand nombre d'explosions: le seul éclatement d'une fenêtre se traduit par une multitude de fragments de verre aux sons tous différents. D'une manière générale, le

« Le jeu vidéo est devenu un produit de très haute technologie »

avec un processeur spécialisé du commerce, autrement dit un matériel très limité en nombre de canaux sonores (64 au maximum) alors que les jeux vidéo ont besoin de beaucoup plus. Pour notre part, nous n'avions pas cette limite hardware car, sur les nouvelles plates-formes de jeux, la philosophie a complètement changé: contrairement aux générations précédentes, qui étaient elles aussi équipées de processeurs spécialisés et devaient donc gérer les mêmes contraintes que celles imposées sur les PC, les nouvelles générations utilisent plusieurs processeurs généralistes, au sein d'une architecture parallèle, ces derniers étant couplés à un processeur spécialisé de rendu audio très simple.

jeu vidéo est devenu un produit de très haute technologie et nous suivons avec attention les publications scientifiques, de même que nous entretenons des relations étroites avec diverses institutions académiques, *via* des bourses de thèse en particulier. » **David Alloza**



EDEN GAMES / ATARI

* Siggraph, acronyme de *Special Interest Group in Graphics*, désigne la conférence internationale réunie chaque année à Los Angeles (Etats-Unis) sur la programmation graphique et les techniques interactives.

* La PlayStation3 est une console de jeu vidéo conçue par Sony Computer Entertainment.

* La Xbox360 est une console de jeu vidéo conçue par Microsoft.